Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

MEASUREMENT OF GROWTH AND EFFICIENCY IN ARITHMETIC (*Continued*)

A YEAR'S PROGRESS

S. A. COURTIS
Home and Day School, Detroit

In a previous article[1] the writer described an attempt made by the Detroit Home and Day School, a private school for girls, to measure the growth of ability in arithmetic through the school by giving the "Stone" tests under identical conditions to every grade from the third to the thirteenth inclusive. Although the experiment was largely a product of chance conditions, the knowledge gained, both of the condition of arithmetic in the school and of the possibilities of the method, seemed to warrant a further experimental study of the same problem by the same means. After due consideration of the various tempting lines of investigation opened, it was decided to try to measure the change produced by a year's regular work under existing conditions in order to have some standard by which to measure the effect of future changes in method. Accordingly, new tests were devised and given at the beginning of the school year, September, 1909, and again at its close, June, 1910. The present article, therefore, will be a brief summary of the results obtained together with some account of the development of the method. As far as possible repetition of previous conclusions will be avoided, and the discussion limited to points of general interest.

In planning the work for the year, the first point to be settled was the scope of the tests themselves. From the work of the previous year it was evident that, as Dr. Stone concludes, ability in arithmetic is no simple thing, but a complex of many separate abilities, and as there was little evidence of transfer from one to another, it seemed necessary to test for each of the separate elements as well as for the general complex. A preliminary study of the specific abilities involved in the Stone tests revealed such a bewildering array that it was decided to limit the new work strictly to the four operations with whole numbers. The specific elements of such work from the standpoint of the classroom were judged to be (1) ability to give a ready response in the case of any of the elementary combinations in the four operations, (2) ability to recognize a situation in a problem as calling for the use of a certain operation, (3) ability to borrow and carry, and (4) the ability to copy correctly the figures of a given example or problem. Accordingly provision was made for the testing of each of these as well as the general abilities involved in working abstract examples and concrete problems.

---

[1] *Elementary School Teacher*, X (No. 4), 177.

To test the first ability, four speed tests were devised, one for each operation, consisting of the 81 fundamental combinations (zero combinations omitted), so arranged that the results only were to be written by the children. Fig. 1 shows a part of the test in addition. These tests will be known hereafter as Speed Tests, Fundamentals. Emphasis was put upon "speed," because the thing to be tested was *readiness* of response. From the fourth grade on every child, if allowed sufficient time, could give the correct answer to every one of the problems in the four tests. Nevertheless, individual children and adults have been found to differ widely in the number of answers written in a given time. The tests measure, therefore, not only *what* is known but *how well* it is known. The time allowance set for these tests was 90 seconds.

SPEED TEST—Addition

Write on this paper as many answers to these simple addition examples as you can in the time allowed.

5	4	2	4	2	2	9	1	2	1	6	7	1	5	1	7	1	2	8	8	3	1	4
7	7	6	4	7	9	9	2	3	8	8	7	4	6	6	8	5	3	8	9	5	3	9

FIG. 1

In the construction of the tests no attention was paid to the question of the relative difficulty of the combinations, or to that of the amount of writing to be done. Consequently the tests are very uneven and the data obtained from their use unreliable. Particularly is this true of the test in subtraction, which from a score of 40 on was very much more difficult than the other tests. This fact must be remembered in considering the scores of the various grades shown in the graphs and tables. However, the use of the tests made evident the necessity for a very careful determination of the value of the different units in each test and the need of extreme caution in making inferences from tests whose components are selected by chance.

Similarly a speed test in reasoning was used to test the ability to recognize a situation as calling for the use of a certain operation. Part of the text is shown in Fig. 2. It consisted of a miscellaneous collection of many simple one-step problems. It will be noted that the children were not asked to work the problems, but to indicate the operation to be used by writing its name in the blank space after an example. The thing tested, however, is again a complex, the ability named being merely its arithmetical phase. Speed in this test was probably determined as much by ability to read quickly and understandingly as by control of knowledge of the operation to be used in the different situations. Here, too, the varying

ARITHMETIC TEST—Reasoning—Speed Test

Do not work the following examples. Read each example through, make up your mind what operation you would use if you were going to work it, then write "add," "subtract," "multiply," or "divide," as the case may be, in the blank space after the example.

1. A pile of fruit contained 89 apples; 27 apples were added to it. How many apples were there in the pile then? ....

2. A girl who had 17 dolls gave away three. How many had she left? ....

3. A thermometer rose 27 degrees in the morning and five times as much in the afternoon. How many degrees did it rise in the afternoon? ....

4. A man worth $60,000 divided his property equally among his 7 children. How much did each one receive? ....

FIG. 2

length of the problems gave units of varying value. The time allowance for this test was two minutes.

The name given to the third ability needs discussion as well as the method by which it was tested. The child is learning to "borrow and carry" all through the 2d, 3d, and 4th grades at least, but a little consideration will show that the ability thus named is not the same in the four operations. In subtraction the number borrowed or carried is always 1, and the procedure is comparatively simple. In addition the number carried varies from 0 to 9 and must be retained a longer time. In multiplication the situation is still more complex owing to the larger numbers involved, the additions that must be performed, and the greater time interval during which the number carried must be retained. Division, the combination of all the other operations, has the difficulties of all, and the added strain due to the constantly varying procedure. There is a common element throughout—the memory factor—but there may well be four abilities instead of one. However, the term "borrowing and carrying" will be used of all as expressing roughly the significance of what is tested in every case.

Provision was made for the testing of this ability by so choosing the figures of the first four examples of the test on Fundamentals (see Fig. 3), that no borrowing or carrying was required. In a comparison of the scores for the first four examples with those for the remaining examples, the increase in the number of mistakes per example is due partly, if not almost wholly, to the effect of this factor. The effect is twofold: actual mistakes in borrowing and carrying may occur or there may be an increase in the mistakes in the combinations used, due to the added strain. Here, however, as in all the tests, the results are not wholly reliable because of the inequalities in the units of which the test is composed.

The ability to copy the figures of the examples and problems correctly was also tested indirectly. A blank piece of paper was supplied, and the instructions directed that examples and problems be copied and worked on this paper. Mistakes in copying were then detected in the scoring.

In addition to these special tests, the two from the previous year were retained but were limited to the four operations (Figs. 3 and 4). In the test on Fundamentals it will be

ARITHMETIC TEST—Fundamentals

Copy the following examples on the sheet of blank paper and work as many as you can in the time allowed. Work them in order, and put down on the paper you hand in all the work you do not do in your head.

1. Add:
a. 2
3
4
1

b. 234
851

2. Subtract:
898456
297125

3. Multiply 231322 by 3.

4. Divide 1243284 by 4.

5. Add:
a. 37
76
89

b. 276
749

6. Subtract:
923
789

7. Multiply 89 by 37.

8. Divide 85666 by 29.

9. Add:
a. 345
684
983
27
405

b. 695797
6926895

10. Subtract:
8749367
5639598

11. Multiply 6789 by 768.

12. Divide 706654443 by 7841.

FIG. 3

ARITHMETIC TEST—Reasoning

Work the following examples on the blank paper, putting down there all the work you do not do in your head.

1. A thermometer stood at 57 degrees at 8 o'clock in the morning. It rose 22 degrees during the morning, and fell 47 degrees during the afternoon. What was the temperature at 6 o'clock?

2. A train-boy made $7.50 on books, magazines, etc., and $8.25 on candy, fruit, etc., each day, for 7 days. What were his total earnings for the week?

3. A man had $2700 in one bank, $1650 in another. If at his death he left his money to his three sons, giving each an equal share, how much did each receive?

4. A man bought 17 old books for 75c apiece and sold them for $1.37 apiece. How much did he make?

FIG. 4

noted that the examples in addition are in two parts. The purpose of this was to test out a conclusion of the previous work, reached also by Dr. Stone, that column addition of several figures is performed as correctly as the shorter two-figure additions which are constantly occurring in multiplications. The two parts of the addition examples, as far as was possible, are composed of such numbers that the additions to be performed are the same. In the test of Reasoning the examples contained the various combinations of operations in two- three- and four-step problems. In both tests the inequalities in the units make comparison of scores of doubtful value. The time allowance for the test in Fundamentals was ten minutes; for Reasoning eight minutes.

During the year it became evident that still another component of the general ability must be tested for—the rate of motor activity in writing figures. Accordingly, after some preliminary work to determine the relative effort involved in writing the different figures, a speed test in copying figures was constructed (Fig. 5). This ability is not to be confused with the ability mentioned above, as in this test all the elements of perception and attention, which in the former test play the major part, have been so reduced to a minimum that the thing tested is largely the rate of motor activity. The time allowance for this test was one minute.

SPEED TEST—Copying Figures

Copy as many of these figures as possible in the time allowed. Write as rapidly as possible, but form the figures as carefully as in working examples:

24967	42976	62947	72964	24967	42967	62974
72946	26974	46927	64972	74926	26947	46972
64927	74962	27946	47962	67924	76942	27964

FIG. 5

All the tests (except the last) were given, under as nearly identical conditions as possible, both in September, during the second week of school, and in June, the week before the close. The tests were printed and a copy of each, together with the necessary blank paper, was given each child. The testing was conducted in the four assembly rooms of the school and the tests were taken by all grades from the 3d to the 13th inclusive. The same person acted as examiner in both tests (except in grades 3, 4, and 13 in June), and the same time allowances, paper, and arrangements held for all from the lowest to the highest. The papers were scored by the teachers and assistants of the mathematics department in accordance with a very definite printed system in which provision was made for the classification of every mistake. In June the scoring was done by one person.

In addition to the above, the speed tests in Fundamentals and in Copying Figures were given in the classrooms by the teachers of the various grades from 4–9 inclusive at intervals during the year—at the end of the first six weeks, before and after the Christmas and Easter vacations, and every week during the last ten weeks of school. The schedule of the tests and the dates at which they were taken was not the same for all grades, but more detailed information on these points can be had, where necessary, from the tables. At least two tests were taken in any one week and often all five.

The purpose of the test at the end of the first six weeks' work was to eliminate the effects of familiarity with the tests and of the recovery in knowledge and ability due to the first weeks of school. During these weeks no specific drill was given in the various tables other than that incidental to the regular work in arithmetic (except in grade 5), but from that time on, the first five minutes (approximately) of each lesson, (three recitations a week) were devoted to specific drill on the combinations. The results as a whole, therefore, give a more or less complete record of the effects of the different phases of school work in arithmetic upon the knowledge of the tables in the four operations.

In scoring the results a record was kept of the work attempted (Ats.) and the part right (Rts.) for each individual. In both these records each combination (figures in the copying test), and each whole example counted one point, the records for the different tests being kept separately. No credit was given for part of an example, the work of the previous year having shown that gross scores by whole examples express the relative ability of grades as well as exact scores where allowance is made for every figure, and at a great saving in time. In the speed tests in Fundamentals, although a record was kept, the results for "Rights" are not given. The errors made were so few as to be negligible. Ignorance of the combinations was shown, not by errors, but by reduction in the number attempted.

Table 1 and Figs. 6 and 7 give for each grade the average gross score in the various tests both in September and in June. A discussion of the distribution of ability within the grades will be found below. The difference between the two averages in each case, or the growth for the year, is also given.

Before taking up the discussion of these results, however, there are one or two points that need special comment.

The time set for the speed tests in Fundamentals proved awkward to control. During the year several mistakes in timing occurred which were corrected at the time either by repeating the test or by discarding the results. From other measurements made of the ability of the 9th grade in subtraction during the year, it is probable that such a mistake was made in the test for subtraction in September in the case of the assembly room of the 9th, 10th, 11th, and 12th grades. Accordingly, the results have been corrected by computing what they would have been in a minute and a half assuming that the actual time allowed was two minutes. The actual values obtained are given by the light dotted lines in the graph. In the case of the test of the 13th grade in Copying Figures, a similar mistake has probably occurred, the corrected and actual results being indicated as before. These mistakes but serve to emphasize the necessity for *identical conditions* in comparative testing where the achievement of one group is to be compared with that of another. They also make it probable that next to the construction of the tests themselves, no other factor within the

TABLE 1

September and **June** Results in All Tests. Grade Averages and Growths

Grade	Addition		Subtraction		Multiplication		Division		Copying Figures	
	Average	Growth	Average	Growth	Average	Growth	Average	Growth	Average	Growth
3	12 **27**	**15**	9 **28**	**19**	. . . .	. .	. . . .	. .	**57**	
4	23 **39**	**16**	19 **37**	**18**	6 **22**	**16**	8 **26**	**18**	**78**	
5	35 **64**	**29**	30 **54**	**24**	24 **56**	**32**	19 **50**	**31**	**107**	
6	43 **63**	**20**	33 **53**	**20**	30 **47**	**17**	26 **48**	**22**	**117**	
7	53 **85**	**32**	36 **64**	**28**	41 **74**	**33**	34 **70**	**36**	**113**	
8	56 **84**	**28**	42 **69**	**27**	49 **72**	**23**	41 **77**	**36**	**116**	
9	67 **81**	**14**	45 (60) **54**	**9**	56 **69**	**13**	46 **64**	**18**	**118**	
10	72 **79**	**7**	47 (63) **54**	**7**	59 **69**	**10**	52 **66**	**14**	**117**	
11	76 **82**	**6**	51 (69) **54**	**3**	62 **68**	**6**	57 **63**	**6**	**116**	
12	74 **79**	**5**	50 (67) **57**	**7**	56 **63**	**7**	53 **62**	**9**	**111**	
13	77 **85**	**8**	57 **62**	**5**	59 **76**	**17**	56 **54**	**−2**	**106** **(141)**	

Grade	Speed Reasoning				Fundamentals				Reasoning			
	Ats.		Rts.		Ats.		Rts.		Ats.		Rts.	
	Ave.	Gth.	Ave.	Gth.	Ave	Gth.	Ave.	Gth.	Ave.	Gth.	Ave.	Gth.
3	2.6	. . .	2.4	. . .	1.8 **3.4**	**1.6**	1.0 **1.6**	**0.6**	. . .	. . .	. . .	. . .
4	5·8 **7.5**	**1.7**	4·2 **4.7**	**0.5**	3·6 **5.5**	**1.9**	2·4 **4.0**	**1.6**	1·1 **4.7**	**3.6**	0·6 **2.6**	**2.0**
5	8·3 **14.6**	**6.3**	7·2 **11.1**	**3.9**	7·0 **9.0**	**2.0**	5·0 **6.3**	**1.3**	4·1 **7.8**	**3.7**	3·4 **5.2**	**1.8**
6	9·2 **13.7**	**4.5**	8.1 **11.5**	**3.4**	8·2 **9.3**	**1.1**	6.2 **6.2**	**0.0**	6·1 **7.7**	**1.6**	3·3 **4.3**	**1.0**
7	9·9 **15.6**	**5.7**	9·0 **13.7**	**4.7**	8·9 **10.9**	**2.0**	6.8 **9.4**	**2.6**	7·2 **8.2**	**1.0**	4·6 **6.3**	**1.7**
8	12·3 **15.4**	**3.1**	11.2 **14.2**	**3.0**	9·9 **11.1**	**1.2**	9.8 **9.1**	**−0.7**	8·0 **8.7**	**0.7**	6·8 **6.7**	**−0.1**
9	11·4 **15.3**	**3.9**	10·4 **13.2**	**2.8**	10·3 **11.0**	**0.7**	7·0 **6.8**	**−0.2**	8·1 **8.8**	**0.7**	6·2 **6.7**	**0.5**
10	14·7 **16.5**	**1.8**	12.6 **14.6**	**2.0**	10·5 **11.2**	**0.7**	7·7 **8.5**	**0.8**	8·8 **8.9**	**0.1**	6·5 **6.8**	**0.3**
11	14·7 **16.7**	**2.0**	13·3 **14.7**	**1.4**	10.8 **11.4**	**0.6**	8.8 **8.2**	**−0.6**	9·0 **8.9**	**−0.1**	6.6 **7.0**	**0.4**
12	15·4 **16.9**	**1.4**	14.2 **15.3**	**1.1**	10.6 **11.2**	**0.6**	8.1 **8.5**	**0.4**	9·4 **9.0**	**−0.4**	7·1 **7.3**	**0.2**
13	14·0 **15.5**	**1.5**	12.8 **13.6**	**0.8**	10·5 **11.2**	**0.7**	8·7 **8.6**	**−0.1**	9·0 **9.1**	**0.1**	6·9 **7.0**	**−0.1**

control of the examiner is so important as the time. The tests for the next year will have a time allowance of even minutes, and some simple mechanical timing device will be adopted.

Interesting as is the record of the year's work to one whose success or failure it discloses, there is space here to discuss only the more general features. Chief among these is the confirmation of the fact shown in the previous work, that there are marked variations of ability, both of different grades in the same test and of the same grade in different tests. The

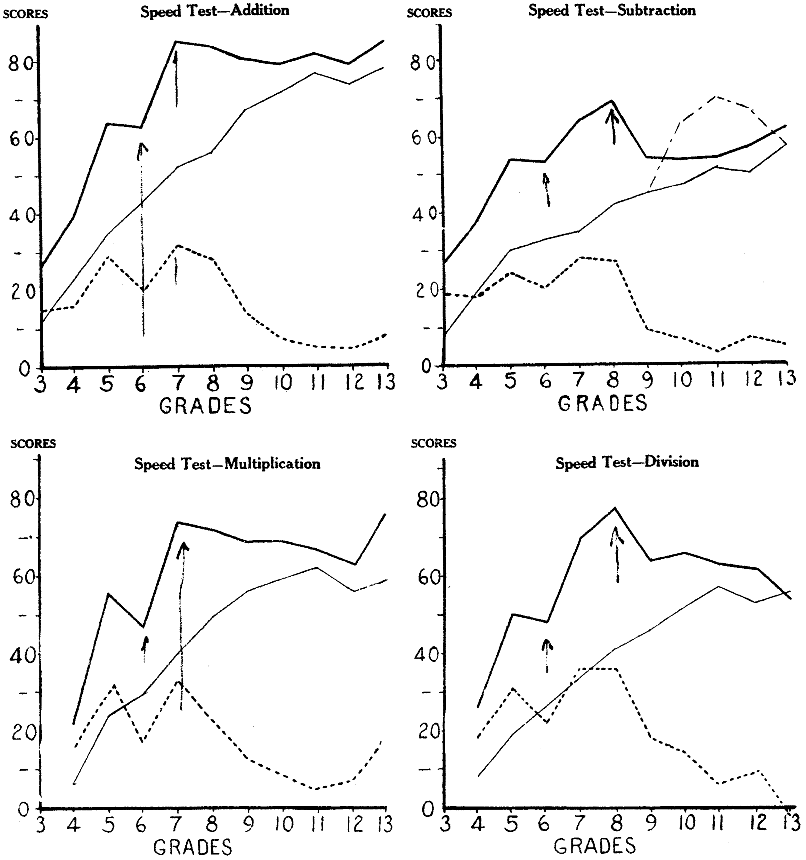

FIG. 6.—September (light lines) and June (heavy lines) grade averages and growths for the year (dotted lines) in tests of knowledge of the tables in each of the four operations.

differences in the amounts of growth for the year show such variations *in the making*. Compare, for instance, the growth of the 8th grade in knowledge of the fundamental combinations in the four operations. The absolute growths are respectively 28, 27, 23, 36. The deviations from the average are – 4 for multiplication and +9 for division, with an extreme difference between the largest and the smallest growths of 13. Trans-

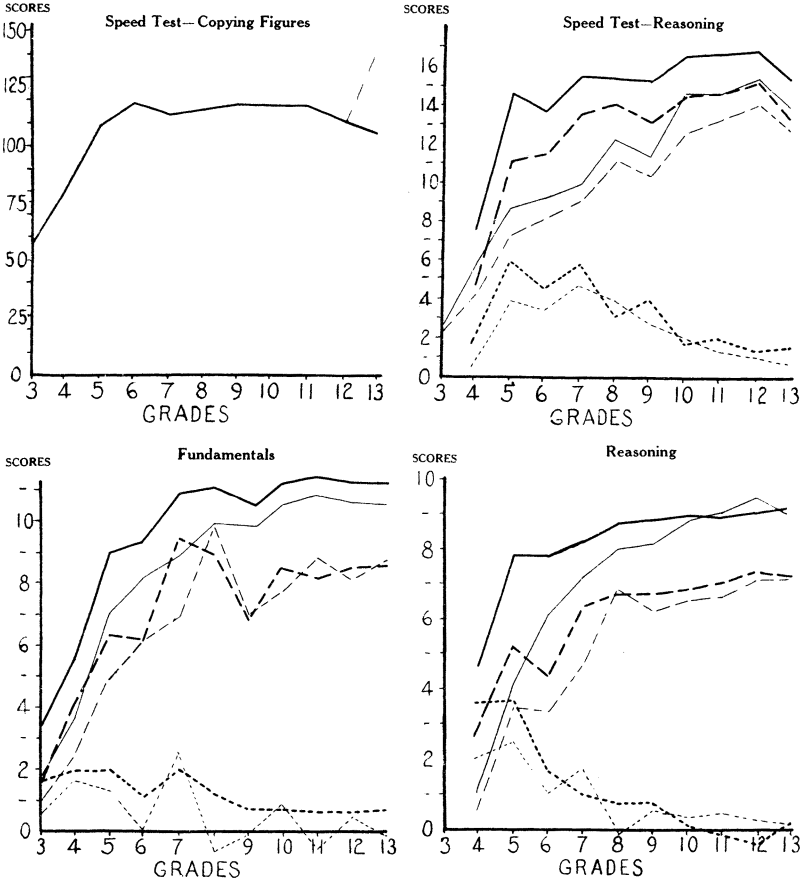

FIG. 7.—September (light lines) and June (heavy lines) grade averages and growths for the year (dotted lines) in the tests shown. The examples attempted are indicated by unbroken lines, those right by broken lines. The growths in examples attempted are shown by the heavy dotted lines, in examples right by the light dotted lines.

lated into weeks this means that nearly six weeks' more work in multiplication and 11 weeks' less work in division was needed to give uniform growths in the four abilities. The percentages of growths are 50, 64, 47, and 88, so that relatively the results are even worse. In September the lowest score was 41 (in division), the highest 56 (in addition), a difference of 15. In June the lowest was 69 (in subtraction), the highest 84 (in addition), the difference, 15, being the same as before. From this point of view, therefore, the year's work has tended slightly to make the scores more uniform. However, it is very evident that the product of "a year's work in arithmetic" is by no means the standard article that our grade system and our courses of study sometimes assume it to be. In other words, each statistical study of educational problems but serves to emphasize the necessity for the constant use of scientific standards and measurements in the schoolroom if there is to be uniformity in the product or efficiency in the teaching effort.

So far as the variations noted above are constant with class after class at the same grade, they point to the operation of causes which ought at least to be accurately and fully determined, if not removed. The defect of the sixth grade is a case in point. This is notoriously a difficult grade. In Dr. Rice's study of the comparative abilities in arithmetic and in spelling of school children in different cities, the scores of the 6th grades are in very many cases below those of the 5th and 7th grades. In the test in arithmetic reported in Scott's *Social Education,* the same drop at the 6th grade is found. Further, the present sixth grade had a high 5th-grade standing last year, while conversely, the poor 6th of last year has made high scores in the present tests. The writer believes the cause to be physical and hopes to be able to prove it so from the data of these tests in connection with the physical measurements taken by the school. Whatever the cause, if it were known, a more efficient adjustment of time and work might be possible.

A similar cause may be operative in grades 8, 9, and 10. The low 9th of the present year was an exceptional 8th-grade class, while the present 10th grade has risen from its previous low position. Further, the present 8th grade shows symptoms of disorganization, as will be seen from the loss in accuracy both in Fundamentals and in Reasoning. The relatively large growth of the 9th grade, greater than that of any other class not studying arithmetic, may be due to a partial recovery of previous loss. So far as these apparent constant variations are really due to general causes, they are likely to affect all subjects alike, and the explanation, when found, may be of general value.

A simple explanation of the eccentric variations shown will at once occur to everyone: namely, that the results of such tests are not to be depended upon. However, repeated measurements during the year, both of grades and of individuals, make it certain that the results obtained do accurately represent the true conditions. A full discussion of this point will be found below.

During the past few years, many radical changes have been made in the methods and texts of the various courses in mathematics through the school, and the results of the testing work are of interest as measures of the effect of such changes. In general, it may be said that the June records of every grade have exceeded the September records of the next higher grade, so that unless the summer vacation causes a great loss of ability at least no harm has been done by the changes. The coming September tests will be needed before the final word can be said on this point. However, as the scores of the 7th and 8th grades in many cases largely exceed those of the high-school grades, the writer is hopeful for the future. On the other hand, it is only fair to call attention to the fact that the September record in all tests is more regular in general form than either the present, or previous, June records. It may be that there is a maximum capacity and an optimum growth for each grade, and that all work in excess of these natural limits is waste effort. The final solution of all such problems must await the cumulative evidence of similar tests through many years.

In this connection a few words must be said in regard to the tendency of the curves for all tests to approach a maximum. This is plainly shown in the graph for the speed test in Copying Figures, where the score for the 6th grade is probably not exceeded by any other grade in the school. The average age of the 6th grade is twelve years, and the general result would seem to mean that complete mastery of such motor control as is essential to speed in writing figures is acquired by this age. Opposed to this is the fact that the range of individual variation in all the classes of the school is very great, and that the range of adult variation is even greater. One thing is certain. In the little remedial work that has been done with individuals up to the present time, no difficulty has been experienced in producing marked increases in scores and no signs of any approach to an upper limit of ability have been seen. The writer believes that the present maxima merely mark the average achievement under present conditions, and that the future will see an increase of from 10 to 50 per cent in the final ability reached. The plans for the coming year include the setting of a standard score for each grade and the effort to bring all below the standard up to grade by special individual work.

The tendency of the scores in the upper grades to approach a maximum has also a bearing on the question of "transfer" of ability, a question of interest to the psychologist and educator alike. The mathematical courses beyond the 8th grade are algebra (grades 9, 10, 13) and geometry (grades 11, 13). No arithmetic is taught in the high-school years. Consequently the growth during the year in these grades is due either to (1) familiarity with the tests and the procedure, i.e., "practice effects"; (2) incidental growth in arithmetic due to its use in the courses in science and mathematics, or in daily life; (3) the effects of maturity, or (4) transfer of ability from the work in other school subjects. It is evident from the small amount of the gains during the year that the part due to "transfer" must be very small, while the slight differences in the scores of these grades make it probable that the "gain" of the year will be offset by the "loss" during the summer. In other words, the present results confirm the conclusion of the previous year that the degree of ability in arithmetic attained by the average child in the grades is not materially increased by this high-school work. Consequently methods which insure completeness and thoroughness in the early training become of very great importance.